



# Primate Extinction Risk and Historical Patterns of Speciation and Extinction in Relation to Body Mass

## Citation

Matthews, Luke J., Christian Arnold, Zarin Machanda, and Charles L. Nunn. 2011. Primate extinction risk and historical patterns of speciation and extinction in relation to body mass. *Proceedings of the Royal Society of London B* 278(1709): 1256-1263.

## Published Version

doi:10.1098/rspb.2010.1489

## Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:8205338>

## Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Open Access Policy Articles, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#OAP>

## Share Your Story

The Harvard community has made this article openly available.  
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23

**Primate extinction risk and historical patterns of speciation and extinction in  
relation to body mass**

Luke J. Matthews<sup>1,\*</sup>, Christian Arnold<sup>1</sup>, Zarin Machanda<sup>1</sup> and Charles L. Nunn<sup>1</sup>

<sup>1</sup> Department of Human Evolutionary Biology  
Harvard University  
11 Divinity Ave.  
Cambridge, MA 02138

\*Author for correspondence ([ljmatth@fas.harvard.edu](mailto:ljmatth@fas.harvard.edu)).

Body mass is thought to influence diversification rates, but previous studies have produced ambiguous results. We investigated patterns of diversification across 100 trees obtained from a new Bayesian inference of primate phylogeny that sampled trees in proportion to their posterior probabilities. First, we used simulations to assess the validity of previous studies that used linear models to investigate the links between IUCN Red List status and body mass. These analyses support the use of linear models for ordinal ranked data on threat status, and phylogenetic generalized linear models revealed a significant positive correlation between current extinction risk and body mass across our tree block. We then investigated historical patterns of speciation and extinction rates using a recently developed maximum likelihood method. Specifically, we predicted that body mass correlates positively with extinction rate because larger-bodied organisms reproduce more slowly, and body mass correlates negatively with speciation rate because smaller-bodied organisms are better able to partition niche space. We failed to find evidence that extinction rates covary with body mass across primate phylogeny. Similarly, speciation rate was generally unrelated to body mass, except in some tests that indicated an increase in speciation rate with increasing body mass. Importantly, we discovered that our data violated a key assumption of sample randomness with respect to body mass. After correcting for this bias, we found no association between diversification rates and mass.

**Keywords:** Primate phylogeny; speciation; extinction; IUCN Red List; body mass; diversification; Bayesian phylogenetics

## 1. INTRODUCTION

Several studies have implicated body size as a risk factor for extinction in mammals (Purvis *et al.* 2000; Cardillo *et al.* 2005; Clauset & Erwin 2008; Davies *et al.* 2008). Large body size is thought to correlate with higher extinction risk through a number of life-history covariates of increased body mass, such as longer generation times and smaller litters (Cardillo *et al.* 2005; Isaac *et al.* 2005). These factors should increase the time needed to recover from stochastic demographic reductions in population size, thus increasing the probability of extinction. To test this hypothesis, studies have quantified extinction risk among living species as a ranked variable derived from the IUCN Red List conservation status categories ([www.iucnredlist.org](http://www.iucnredlist.org), Purvis *et al.* 2000; Cardillo *et al.* 2005; Davies *et al.* 2008). In mammals, Cardillo *et al.* (2005) found that “intrinsic” biological life history traits influenced extinction risk only when species were >3 kg in mass. Below 3 kg mass, the primary determinants of extinction risk were “extrinsic” factors that are not genetically heritable, such as geographic range size and nearby human population density (Cardillo *et al.* 2005). Studies also have failed to detect a consistent association between diversification rate and body mass across mammalian clades (Gittleman & Purvis 2000; Isaac *et al.* 2005).

Understanding the predictors of present day extinction risk, as reflected by the IUCN Red List, has clear relevance given the pressing need for primate conservation (Cowlshaw & Dunbar 2000). For example, the great apes are all large-bodied and critically endangered (IUCN Red List). An important question is whether large bodied primate species experience higher extinction rates irrespective of human activities. If so, then large bodied primates may be especially vulnerable to anthropogenic drivers of

extinction, such as reductions in habitat or climate change. It is possible to infer speciation and extinction rates from a dated phylogeny because higher rates of extinction relative to speciation should produce longer internal branches on a tree with an apparent burst of diversification close to the tips (Nee *et al.* 1994; Nee 2001). Recent advances provide a way to integrate the study of speciation and extinction that is dependent upon another trait (Maddison *et al.* 2007), such as body mass. Previous studies have investigated patterns of diversification more generally in primates (Purvis *et al.* 1995; Gittleman & Purvis 1998; Chan & Moore 2002; Moore *et al.* 2004; Paradis 2005; Freckleton *et al.* 2008) using an older inference of primate phylogeny (Purvis 1995).

Using a newly inferred primate phylogeny that enabled us to incorporate phylogenetic uncertainty, we investigated the links between body mass, extinction risk and diversification rates. First, we tested for an association between body mass and present-day extinction risk categories obtained from the IUCN Red List. As with previous studies, we predicted that larger-bodied primates are at greater risk of extinction (Purvis *et al.* 2000; Cardillo *et al.* 2005; Davies *et al.* 2008). In our analyses, we used methods that incorporate phylogenetic uncertainty, that better model character evolution on the tree, and that test for spurious results arising from the ordinal (rather than integer) measurement of extinction threat categories.

Second, we predicted that primate lineages characterized by greater body mass have experienced higher extinction rates throughout evolutionary history. For this, we applied a new method to estimate extinction rates on a phylogenetic tree in relation to a biological characteristic (Maddison *et al.* 2007), again running the test in a way that incorporated phylogenetic uncertainty.

Lastly, we turned our attention to speciation by predicting that lineages with smaller body mass have higher speciation rates. We made this prediction based on several studies showing a trend for smaller bodied mammalian clades to be more diverse, possibly because they experience vicariance events more frequently or they are able to partition the environment into more niches (Maurer *et al.* 1992; Gittleman & Purvis 1998; Gardezi & da Silva 1999; Purvis *et al.* 2000; Cardillo *et al.* 2005; Isaac *et al.* 2005).

## 2. METHODS

### (a) *Tree Inference*

As our hypotheses pertained to the order Primates as a whole, we needed a phylogeny that included as many species as possible. To this end, we used the trees available from the *10kTrees* project, Version 1 (<http://10ktrees.fas.harvard.edu/>). The website provides extensive documentation on the tree inference, a graphical interface for downloading trees, and a number of visualizations of the trees. Details regarding tree inference are available in Arnold *et al.* (2010). Our use of Bayesian tree inference enabled us to deal with phylogenetic uncertainty by running comparative tests on multiple trees saved from the Markov chain. Allowing for such topological and branch length uncertainty is important because the phylogeny used can affect the conclusions that are drawn from a comparative analysis (Lutzoni *et al.* 2001, Pagel & Lutzoni 2002).

We used two sets of trees from the Bayesian analysis: a sample of 100 trees distributed evenly along the post-burnin Markov chain and a consensus tree of all nodes with clade credibility support greater than 0.5. We dated all trees prior to comparative analysis by using seven fossil calibration points employed by previous phylogenetic

studies (Table 1, Seiffert *et al.* 2003; Yang & Yoder 2003; Yoder & Yang 2004; Godinot 2006; Hodgson *et al.* 2009). We conducted molecular dating with the software r8s (Sanderson 2002) using the penalized likelihood algorithm with a smoothing parameter of 100, chosen because this value best recovered dates inferred from phylogenetic analyses of smaller taxonomic samples but with more extensive sequence data (Yang & Yoder 2003; Yoder & Yang 2004; Hodgson *et al.* 2009).

#### **(b) *Body mass and IUCN Red List data***

Body mass data were obtained from Smith & Jungers (1997). We calculated the mean female body mass across study sites for our analysis. In this manner, we obtained body mass data on 160 species that could be matched to our phylogeny through translation via the taxonomy of Wilson & Reader (2005). Analyses used the natural log of female body mass. We obtained IUCN conservation status for the species on our phylogeny from the IUCN Red List website ([www.iucnredlist.org](http://www.iucnredlist.org)). From each IUCN category we constructed an ordinal variable with higher ranks corresponding to greater extinction threat. Specifically, ranked from smallest to highest, we used the following categories of threat: least concern, near threatened, vulnerable, endangered and critically endangered. Six species labeled “data deficient” in the IUCN Red List were removed from the analysis, reducing the sample size to 154 species. The highest rank was assigned to “critically endangered” rather than “extinct in the wild” because our study included only extant species as data points.

137 **(c) *Comparative analyses***

138 We tested for an association between body size and IUCN extinction risk through a  
139 phylogenetic generalized least squares (PGLS) model applied across our Bayesian tree  
140 block. This analysis of the primate data replicates previous independent contrasts  
141 analyses across mammals (e.g. Cardillo et al. 2005; Davies et al. 2008). However, PGLS  
142 implemented in the program BayesTraits (Pagel & Meade 2007) enabled us to  
143 incorporate Bayesian estimation of the branch scaling parameter  $\lambda$  (Pagel 1999) across a  
144 tree block, which should improve upon contrasts-based approaches. In particular, we  
145 suspected that Brownian motion along the branch lengths in our block of 100 trees might  
146 not reflect evolutionary change in a character such as threat status. Brownian motion  
147 along branch lengths is an assumption of independent contrasts. The scaling parameter  
148 lambda adjusts the internal branch lengths with a length multiplier such that the data meet  
149 the assumption of Brownian motion. By systematically searching through many possible  
150 values for lambda, the program locates the value that makes the data most likely and thus  
151 best accommodates the assumption of Brownian motion. In primates, Purvis et al. (2005)  
152 showed that IUCN rank alone exhibited a lambda equal to 0.77 that excluded both zero  
153 and one (95% CI 0.51-0.90). Thus, we considered it important to incorporate lambda in  
154 our PGLS, and especially to do so in a way that incorporates uncertainty in phylogenetic  
155 relationships and branch lengths.

156 The use of PGLS or independent contrasts treats IUCN extinction risk as a  
157 continuous variable (Felsenstein 1985; Garland *et al.* 1992). Counter to this assumption,  
158 extinction risk codes in the IUCN Red List are not continuously varying. Instead,  
159 extinction risk is an ordinal variable in which ranks likely vary in the amount of



160 difference in the actual underlying extinction risk (Purvis et al. 2005). For example, the  
161 true (quantitative) difference in extinction risk may differ between categories of near-  
162 threatened and vulnerable, as compared to endangered and critically endangered.  
163 Treating ordinal variables as continuous can produce elevated Type 1 error rates because  
164 such treatment applies arithmetic operations that do not preserve the variance structure of  
165 the original ordinal ranks (Stevens 1946; Siegel 1956). The problem occurs specifically  
166 when the ordinal ranks are separated by unequal distances along the underlying  
167 continuous variable that they measure, a point acknowledged by Purvis et al. (2005) and  
168 given in the example above.

169       We used computer simulations to assess whether the treatment of IUCN threat  
170 categories as continuously varying may have introduced error into this study and previous  
171 studies (Cardillo et al. 2005; Purvis et al. 2005; Davies et al. 2008). Specifically, we  
172 tested whether treating threat status as continuously varying results in elevated Type 1  
173 error rates. We conducted 1000 simulations of two uncorrelated continuous characters on  
174 our consensus tree (function “sim.char” in the R package “geiger,” Harmon *et al.* 2009).  
175 Characters evolved randomly with a constant accumulation of variance (set at 1.0 per unit  
176 branch length) and an initial state of zero. We then rescored one character from each pair  
177 into a set of ordinal ranks with the same number of species in each state as was observed  
178 for each corresponding IUCN threat status in our data set. The lowest continuous values  
179 of the simulated character thus became rank “0,” while the highest became rank “4,” with  
180 other ranks derived from intermediate values and all in matching proportion to the  
181 observed frequency of each rank in the IUCN data. This rendered one character a true  
182 continuous distribution, like body size, but the other character had been rescored into a

set of ordinal ranks with different real distances between the means of each rank. We then conducted 1000 PGLS tests to test for a significant association of the two characters, one of which was ordinal and one continuous. Significant associations were counted as Type 1 errors because the characters evolved independently during the simulations. An alternative approach to deal with ordinal IUCN data would be to include additional parameters that model the ordinal nature of the dependant variable. While this procedure in principle can be implemented (Hadfield & Nakagawa 2010), given the prominent prior research that treated IUCN as continuous, we prefer to use simulations to test this approach more generally.

To test for effects of body size on the actual speciation and extinction rates experienced by different lineages on the primate tree, we employed the “binary-state speciation and extinction” (BiSSE) test (Maddison *et al.* 2007), as implemented in the R package diversitree (FitzJohn 2009; FitzJohn *et al.* 2009). This procedure uses likelihood methods to test a six-parameter model of speciation, extinction, and trait evolution. Following the procedure recommended in Maddison *et al.* (2007), we constructed five parameter and six parameter models to test body-size dependant speciation and extinction rates, which we investigated separately. For each state of a binary character, BiSSE can model a rate of speciation, extinction, and character state transitions. Thus, six rates are possible for the most complex BiSSE model. The five-parameter models constrained the speciation (or extinction) rates to be equal in the two body size categories, while the six-parameter models allowed speciation (or extinction) rates to vary for different body sizes. Because these models were nested, we assessed statistical significance using likelihood ratio tests. We conducted all BiSSE analyses on our dated consensus tree.

206           To use BiSSE with our measures of body mass, we binned species into categories  
207 of “small” and “large” body mass. As it is not immediately clear what cutoff should be  
208 used for these categories, we initially tested our hypotheses with three different cutoffs  
209 derived from the primatological literature. The first break point, 500 g, reflects the point  
210 above which primate species are not strictly faunivorous (i.e., Kay’s threshold, Kay  
211 1984). We thought this widely accepted energetic constraint might reflect the life history  
212 variables that underlie previously observed associations between body mass and  
213 extinction risk. The second break point, 984 g, was the phylogenetic mean body mass at  
214 the root of the dated consensus tree, as inferred through squared change parsimony  
215 reconstruction (Maddison 1991) of log body mass values in Mesquite version 2.6  
216 (Maddison & Maddison 2006). The third break point of 3000 g came from a study of  
217 extinction risk across mammals (Cardillo *et al.* 2005), in which the authors found that  
218 biologically intrinsic life history traits only influenced extinction risk above a body mass  
219 of 3000 g. We ran these analyses of the three *a priori* break points across the 100 dated  
220 trees from the Bayesian tree search.

221           To assess the sensitivity of our results to the break point used for binning species,  
222 we conducted an analysis with a sliding break point that divided the body mass data into  
223 small and large categories along 20 intervals of 0.2 log units of body mass. We  
224 quantified the precision of the sliding break point estimates by sampling the six-  
225 parameter model with a Monte-Carlo Markov Chain (MCMC) that we initialized with the  
226 maximum likelihood estimates for parameter values (R package diversitree, FitzJohn  
227 2009; FitzJohn *et al.* 2009). We used only the dated consensus tree for the sliding break  
228 point analysis.

The implementation of BiSSE in diversitree allows one to specify the degree of species sampling employed in the study. Doing so is important because methods to test trait-dependent extinction are biased by incomplete sampling of the study group (Nee et al. 1994; FitzJohn et al. 2009). The diversitree package includes modified likelihood equations to account for this bias (FitzJohn et al. 2009). Under the Wilson and Reeder (2005) taxonomy, which served as the taxonomy for the present study, we sampled 42% of extant Primates. We used this value as our sampling percentage in all BiSSE analyses.

The BiSSE correction for incomplete sampling is statistically valid only when the species have been sampled randomly. Two issues are relevant here: random sampling of species from the phylogeny, and random sampling with respect to the character hypothesized to affect speciation and extinction rates.

To assess the randomness of our species sample with respect to phylogeny, we conducted a G-test of proportions of species from each genera within our observed sample (Sokal & Rohlf 1995; R script). The G-test uses maximum likelihood techniques to assess whether an observed proportion of species within genera can be viewed as a random sample of species given their known frequencies from the complete taxonomy (Wilson & Reeder 2005). We conducted this test at the generic level because, based on our primate phylogeny (Arnold *et al.* 2010), generic classifications reflect monophyletic clades.

To test whether sampling was random with respect to body mass, we calculated the deviation of the observed number of species sampled from a genus compared to the number expected to be sampled given the number of species in each genus in the complete taxonomy. Thus, a positive deviation indicated that more species were sampled

from a genus than expected under random sampling, while a negative number indicated fewer were sampled than expected. We then conducted a phylogenetic generalized least squares (PGLS) test for an association between body mass and sampling deviation using our dated consensus tree to control for phylogenetic non-independence.

We then retested our hypothesis regarding speciation and body size using the diversification test of Freckleton et al. (2008). This test is simply a PGLS of a character trait as a dependant variable with the number of nodes from the tree root to each tip as the independent variable. If a particular character state, such as body size, is associated positively with diversification rates, then a positive association should exist between the character value and lineages that exhibit more nodes. The advantage of this test is that, because it is conducted through a PGLS framework, we were able to include the deviations from the randomly expected sampling as an independent variable; in other words, we could investigate the effects of body mass on diversification while controlling for body mass-related biased sampling if it exists, even if not statistically significant in the above tests. Another advantage of this test is it treats body size as a continuous variable, whereas the use of BiSSE required us to bin body size into a set of binary categories. The disadvantage of the Freckleton et al. (2008) test is that it investigates the association between net diversification rate and body size, rather than specifically investigating speciation and extinction separately (as in BiSSE). We used the dated consensus tree for this analysis.

### **3. RESULTS**

#### **(a) *Extinction risk and body mass***

In our dataset, 68 species were categorized as “Least Concern”, 9 as “Near Threatened”, 30 as “Vulnerable”, 40 as “Endangered”, and 7 as “Critically Endangered” (Supplemental Table 1). We found a significant and positive association between IUCN extinction risk and body mass while in a PGLS analysis across our dated Bayesian tree block ( $N = 154$ ,  $\beta = 0.40$ ,  $p = 0.004$ ). Additionally, we found that lambda differed from zero and one ( $\lambda = 0.78$ , 95% credible interval  $\lambda = 0.63 - 0.89$ ). This indicates significant phylogenetic signal and thus highlights the importance of controlling for phylogeny, but further shows that a non-Brownian component may contribute to variation in residual values. In our 1000 simulations that tested the effect of ordinal ranking for IUCN, only 5% of statistical tests were deemed significant at the 0.05 nominal level.

#### **(b) *Extinction rate and body mass***

In our BiSSE analyses of 160 primate species, likelihood ratio tests supported no significant relationship between log body mass and historical extinction rate either at our *a priori* break points or in our sliding break point analysis (Figure 1a). Similarly, across the 100 trees, extinction rates at the *a priori* break points were significantly different relative to body mass for only 1 of the 100 trees sampled from our tree block. This was true even though extinction rates were estimated to be different from zero in many cases (Figure 1a). Several papers have indicated that methods for inferring trait dependant extinction may suffer from low statistical power (Maddison *et al.* 2007; Kubo & Iwasa 1995; Purvis *et al.* 1995; Gardezi & da Silva 1999; Paradis 2004), a point to which we return in the Discussion.

298 **(c) *Speciation rate and body mass***

299 In contrast to the results from the extinction analysis, we found significant speciation rate  
300 differences at the Kay's threshold *a priori* break point on our dated consensus phylogeny  
301 and across 95 of the 100 trees (for the consensus tree,  $p = 0.016$ ). Differences in  
302 speciation rate at the phylogenetic mean and 3 kg break points were not significant on the  
303 consensus tree or on any of the trees from the tree block. Importantly, the result for  
304 Kay's Threshold was opposite to our prediction, with higher estimated speciation rates in  
305 larger-bodied species (Figure 1b). This pattern seemed to apply across most of the  
306 variation in primate body mass, but not at the largest body mass categories (Figure 1b).  
307 However, MCMC sampling of the speciation estimates showed that 95% credible  
308 intervals overlapped across most of the distribution (Figure 1b).

309 In testing for speciation and extinction rates, BiSSE also estimated two rates of  
310 transition (small to large; large to small) in all analyses. The transition rate estimates  
311 showed overlapping credible intervals across almost the entire range of sliding break  
312 points (Supplemental Figure 1). This indicates rates of body size increase and decrease  
313 are generally equivalent across the range of primate body sizes.

314 The BiSSE results are not compromised by nonrandom sampling of species in our  
315 dataset with respect to phylogeny. The G-test for independence showed that the observed  
316 sampling of species within genera was within the variation expected from random  
317 sampling of taxa ( $p = 0.147$ ). However, the observed sample of species was biased  
318 toward greater sampling of species with larger body mass in a PGLS test ( $\beta = 0.80$ ,  $p =$   
319  $0.037$ ). Thus, biased sampling towards species with larger body mass may have  
320 influenced the BiSSE results; specifically, the denser sampling of the larger species may

have biased the test to estimate a higher rate of speciation because these lineages would have occurred more commonly on the tree representing the sample of species.

When we controlled for the observed sampling deviations from the random sampling expectation, our PGLS test of diversification effects on body size did not find any association between these variables ( $\beta = 0.00$ ,  $p = 0.71$ ). When we removed the sampling variable from the analysis, however, the PGLS obtained a trend more consistent with the results of the BiSSE analysis ( $\beta = 0.08$ ,  $p = 0.081$ ). Although this PGLS test cannot technically distinguish between body-mass related speciation and extinction effects, we would expect to find a positive association between diversification and body mass if speciation rate correlates positively with body mass and extinction rates do not (as suggested by the above tests). Additionally, simulations by Freckleton et al. (2008) showed that their test was much more sensitive to speciation effects than to extinction effects. Thus, given the much lower power of the test to detect an effect of extinction, any positive result would more likely result from speciation effects.

#### 4. DISCUSSION

We tested a series of predictions involving the links among body mass, phylogenetic diversification, and current extinction risk in primates. With regard to current patterns of extinction, we found that larger bodied species experience higher extinction risk. This result replicates previous findings, but for the first time allows for phylogenetic uncertainty and uses the scaling parameter  $\lambda$  to control for the non-Brownian distribution of variation in IUCN threat status categories (Cardillo *et al.* 2005; Clauset & Erwin 2008; Davies *et al.* 2008). We also tested, for the first time, whether the ordinal nature of



IUCN threat status categories impacts the statistical performance of independent contrasts and PGLS analyses that use it as a dependent variable. Our 1000 phylogenetic simulations show the ordinal coding itself does not produce elevated type 1 error in our primate dataset.

Despite finding a strong association between threat categories and body mass in primates, we failed to find evidence for an historical association between body mass and extinction rates. Instead, in a limited number of tests, we found that speciation rates may increase with body mass, which was opposite to our predictions. These analyses point to an association between larger body mass and higher speciation rate, which contrasts with most predictions, including ours, that speciation rates covary negatively with body mass (see also Liow *et al.* 2008). After controlling for biased sampling of species in a PGLS model, however, we found no association between diversification rate – a function of speciation and extinction – and body mass.

Previous studies found that intrinsic biological variables can explain a substantial proportion of variation in risk status (Purvis *et al.* 2000; Cardillo *et al.* 2005). These previous statistical models found that high risk variables, including large body mass, high trophic level, small geographic range, and slow life history, explain threat status to a substantial degree and independently of anthropogenic effects (Purvis *et al.* 2000; Cardillo *et al.* 2005; Purvis 2008). Thus, present day extinction risk status should be related to the historically experienced extinction rates of lineages.

We found it surprising that current patterns of extinction risk covaried with body mass in our study and previous studies, while estimates of extinction rates across primate phylogeny failed to show this effect. It could be that higher threat categories among

large-bodied primates have resulted only from anthropogenic effects in the present. Alternatively, the tests may have failed to detect differences because available methods to estimate historical extinction rates suffer from low statistical power (Maddison *et al.* 2007; see also Nee *et al.* 1992; Kubo & Iwasa 1995; Purvis *et al.* 1995; Gardezi & da Silva; 1999; Paradis 2004; Rabosky 2010). Using simulations, for example, Maddison *et al.* (2007) showed that the power to detect variation in speciation rates is relatively high (power for speciation in simulations of 500 species was ~58%), as compared to the power of detecting variation in extinction rates (~21%). Rabosky (2010) recently showed that, in addition to low power, current methods cannot accurately estimate extinction rates when these rates vary across the tree. His analyses were conducted with complete sampling and phylogenetic information of every extant species. Given the fossil evidence for a plethora of extinct primate species, we agree with Rabosky (2010) that future analyses should incorporate fossil information as a means to increase power and to estimate extinction rates more accurately (see also Purvis 2008).

Our study also offers a cautionary tale about the importance of testing the assumptions of methods for studying speciation and extinction rates. When using a new method to account for incomplete sampling of taxa, we found some support for an association between speciation and body mass. However, our data violated another assumption of these methods, namely that sampling of species is random with respect to body mass. When we controlled for non-random sampling, the association between diversification rates and body mass became non-significant. Thus, we offer a valuable empirical example of shortcomings to these methods that complement previous simulation studies, and we provide a procedure for investigating biased sampling for

others to use in future studies. The potential effects of nonrandom sampling are a serious concern for previous studies of trait dependent diversification (Purvis *et al.* 1995; Paradis 2005; Freckleton *et al.* 2008; Nakagawa & Freckleton 2008).

We began this study with the prediction that populations of smaller bodied animals may speciate more quickly, as they experience vicariance events more frequently or they are able to better partition the environment into more niches (Gardezi & da Silva 1999). Previous studies provided hints that speciation rate could increase with body mass in primates, rather than showing a strict negative correlation with body mass. For example, Purvis *et al.* (1995) found that the family Cercopithecidae (Old World monkeys) experienced larger speciation rates than did other primate lineages. While their study did not assess how speciation covaried with characteristics such as body mass, our results are consistent with their findings given that the cercopithecids are generally larger bodied than extant platyrrhine and strepsirrhine primates that more frequently fall below Kay's Threshold. The cercopithecids also are typically smaller than the (relatively) species depauperate hominoids. These clade-specific patterns could be tested more rigorously with the new statistical software MEDUSA (Alfaro *et al.* 2009). Our results are consistent with Freckleton *et al.*'s (2008) findings of a significant correlation between body mass and diversification rate. The results imply that some degree of species selection on body mass may have occurred in primates, in that the influence of body mass on cladogenesis explains the distribution of body mass even after allowing for anagenic change as it is incorporated into the model (Supplemental Figure 1; Vrba 1984; Purvis *et al.* 1995; Maddison *et al.* 2007).

412           In summary, we used new methods to incorporate phylogenetic uncertainty and to  
413    assess how body mass influences primate diversification. Our results add to a growing  
414    body of evidence that larger bodied animals are more susceptible to extinction. The lack  
415    of significance for historical patterns of extinction may indicate that methods to detect  
416    extinction rates from extant species are severely compromised in statistical power (Kubo  
417    & Iwasa 1995; Purvis *et al.* 1995; Gardezi & da Silva 1999; Paradis 2004; Maddison *et*  
418    *al.* 2007; Rabosky 2010). Intriguingly, we found some evidence that speciation rates  
419    appear to increase with body mass in primates (see also Liow 2008). It will be interesting  
420    to see if similar patterns occur in other groups of vertebrates and, if so, how and why the  
421    effects of body size on speciation vary among clades.

424 Acknowledgments

425

426 This research was supported by Harvard University. We thank Daniel Lieberman for his  
427 helpful discussions regarding diversification mechanisms, Richard Wrangham's lab  
428 group for helpful comments, and two anonymous referees and Andy Purvis for valuable  
429 suggestions during the review process.

430

## REFERENCES

- Aitchison, J. 1983 Principal Component Analysis of Compositional Data. *Biometrika* **70**, 57-65.
- Alfaro, M. E., Santini, F., Brock, C., Alamillo, H., Dornburg, A., Rabosky, D. L., Carnevale, G., & Harmon, L. J. 2009 Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. *Proceedings of the National Academy of Sciences, U.S.A.* **106**, 13410-13414.
- Alroy, J. 1998 Cope's rule and the dynamics of body mass evolution in North American fossil mammals. *Science* **280**, 731-734.
- Arnold, C., Matthews, L. J. & Nunn, C. L. 2010 The 10kTrees Website: A New Online Resource for Primate Phylogeny. *Evolutionary Anthropology* **19**, 114-118.
- Cardillo, M., Mace, G. M., Jones, K. E., Bielby, J., Bininda-Emonds, O. R. P., Sechrest, W., Orme, C. D. L. & Purvis, A. 2005 Multiple causes of high extinction risk in large mammal species, vol. 309, pp. 1239-1241: American Association for the Advancement of Science.

453 Chan, K. M. A. & Moore, B. R. 2002 Whole-tree methods for detecting differential  
 454 diversification rates. *Systematic Biology* **51**, 855-865.  
 455  
 456 Clauset, A. & Erwin, D. H. 2008 The evolution and distribution of species body size.  
 457 *Science* **321**, 399-401.  
 458  
 459 Cooper, N. & Purvis, A. 2009 What factors shape rates of phenotypic evolution? A  
 460 comparative study of cranial morphology of four mammalian clades. *Journal of*  
 461 *Evolutionary Biology* **22**, 1024-1035.  
 462  
 463 Cowlshaw, G. & Dunbar, R. I. M. 2000 *Primate Conservation Biology*. Chicago:  
 464 University of Chicago Press.  
 465  
 466 Davies, T. J., Fritz, S. A., Grenyer, R., Orme, C. D. L., Bielby, J., Bininda-Emonds, O. R.  
 467 P., Cardillo, M., Jones, K. E., Gittleman, J. L., Mace, G. M. & Purvis, A. 2008  
 468 Phylogenetic trees and the future of mammalian biodiversity. *Proceedings of the National*  
 469 *Academy of Sciences of the United States of America* **105**, 11556-11563.  
 470  
 471 Felsenstein, J. 1985 Phylogenies and the comparative method. *American Naturalist* **125**,  
 472 1-15.  
 473  
 474 FitzJohn, R. G. 2009. diversitree: comparative phylogenetic tests of diversification. R  
 475 package version 0.2-1. <http://www.zoology.ubc.ca/prog/diversitree>

476

477 FitzJohn, R. G., Maddison, W. P. & Otto, S. P. 2009 Estimating Trait-Dependent  
478 Speciation and Extinction Rates from Incompletely Resolved Phylogenies. *Systematic*  
479 *Biology* **58**, 595-611.

480

481 Freckleton, R. P., Phillimore, A. B. & Pagel, M. 2008 Relating traits to diversification: A  
482 simple test. *American Naturalist* **172**, 102-115.

483

484 Gardezi, T. & da Silva, J. 1999 Diversity in relation to body size in mammals: A  
485 comparative study. *American Naturalist* **153**, 110-123.

486

487 Garland, T., Midford, P. E. & Ives, A. R. 1999 An introduction to phylogenetically based  
488 statistical methods, with a new method for confidence intervals on ancestral values.  
489 *American Zoologist* **39**, 374-388.

490

491 Gittleman, J. L. & Purvis, A. 1998 Body size and species-richness in carnivores and  
492 primates. *Proceedings of the Royal Society of London Series B-Biological Sciences* **265**,  
493 113-119.

494

495 Harmon, L., Weir, J., Brock, C., Glor, R., Challenger, W., & Hunt, G. 2009 geiger:  
496 Analysis of evolutionary diversification. R package version 1.3-1. [http://CRAN.R-](http://CRAN.R-project.org/package=geiger)  
497 [project.org/package=geiger](http://CRAN.R-project.org/package=geiger)

498



499 Hodgson, J. A., Sterner, K. N., Matthews, L. J., Burrell, A. S., Jani, R. A., Raaum, R. L.,  
 500 Stewart, C. B. & Disotell, T. R. 2009 Successive radiations, not stasis, in the South  
 501 American primate fauna. *Proceedings of the National Academy of Sciences of the United*  
 502 *States of America* **106**, 5534-5539.  
 503  
 504 Huelsenbeck, J. P., Rannala, B. & Masly, J. P. 2000 Accommodating phylogenetic  
 505 uncertainty in evolutionary studies. *Science* **288**, 2349-2350.  
 506  
 507 Huelsenbeck, J. P. & Ronquist, F. 2001 MrBayes: Bayesian inference of phylogeny.  
 508 *Bioinformatics* **17**, 754-755.  
 509  
 510 Isaac, N. J. B., Jones, K. E., Gittleman, J. L. & Purvis, A. 2005 Correlates of species  
 511 richness in mammals: Body size, life history, and ecology. *American Naturalist* **165**, 600-  
 512 607.  
 513  
 514 Jackson, D. A. 1997 Compositional data in community ecology: The paradigm or peril of  
 515 proportions? *Ecology* **78**, 929-940.  
 516  
 517 Kay, R. F. 1984 On the use of anatomical features to infer foraging behavior in extinct  
 518 primates. In *Adaptations for Foraging in Nonhuman Primates* (ed. P. Rodman & J. Cant).  
 519 New York: Columbia University Press.  
 520

521 Kaufman, L. & Rousseeuw, P. J. 1990 Finding groups in data : an introduction to cluster  
522 analysis. *Wiley series in probability and mathematical statistics. Applied probability and*  
523 *statistics*, , xiv, 342 p.

524

525 Kubo, T. & Iwasa, Y. 1995 Inferring the Rates of Branching and Extinction from  
526 Molecular Phylogenies. *Evolution* **49**, 694-704.

527

528 Legrende, P., Lapointe, F. & Casgrain, P. 1994. Modeling brain evolution from behavior:  
529 a permutational regression approach. *Evolution* **48**, 1487-1499.

530

531 Lehman, S. M. 2007 Ecological and phylogenetic correlates to body size in the Indriidae.  
532 *International Journal of Primatology* **28**, 183-210.

533

534 Liow, L. J, Fortelius, M., Bingham, E., Lintulaakso, K., Mannila, H., Flynn, L. &  
535 Stenseth, N. C. 2008 Higher origination and extinction rates in larger mammals.  
536 *Proceedings of the National Academy of Sciences* **105**, 6097-6102.

537

538 Losos, J. B., Leal, M., Glor, R. E., de Queiroz, K., Hertz, P. E., Schettino, L. R., Lara, A.  
539 C., Jackman, T. R. & Larson, A. 2003 Niche lability in the evolution of a Caribbean  
540 lizard community. *Nature* **424**, 542-545.

541

542 Lutzoni, F., Pagel, M. & Reeb, V. 2001 Major fungal lineages are derived from lichen  
543 symbiotic ancestors. *Nature* **411**, 937-940.

544

545 Maddison, W. 1991 Squared-change parsimony reconstructions of ancestral states for  
 546 continuous-valued characters on a phylogenetic tree. *Systematic Zoology* **40**, 304-314.

547

548 Maddison, W. P., Midford, P. E. & Otto, S. P. 2007 Estimating a binary character's effect  
 549 on speciation and extinction. *Systematic Biology* **56**, 701-710.

550

551 Maddison, W. P. & Maddison, D. R. 2009 Mesquite: a modular system for evolutionary  
 552 analysis. Version 2.6 <http://mesquiteproject.org>

553

554 Maurer, B. A., Brown, J. H. & Rusler, R. D. 1992 The Micro and Macro in Body Size  
 555 Evolution. *Evolution* **46**, 939-953.

556

557 Moore, B. R., Chan, K. M. A. & Donoghue, M. J. 2004 Detecting diversification rate  
 558 variation in supertrees. *Phylogenetic Supertrees: Combining Information to Reveal the*  
 559 *Tree of Life*, 487–533.

560

561 Nakagawa, S & Freckleton, R. P. 2008 Missing inaction: the dangers of ignoring missing  
 562 data. *Trends in Ecology & Evolution* **23**, 592-596.

563

564 Nee, S. 2001 Inferring speciation rates from phylogenies. *Evolution* **55**, 661-668.

565

566 Nee, S., Mooers, A. O. & Harvey, P. H. 1992 Tempo and Mode of Evolution Revealed  
567 from Molecular Phylogenies. *Proceedings of the National Academy of Sciences of the*  
568 *United States of America* **89**, 8322-8326.

569

570 Nee, S., Holmes, E. C., May, R. M. & Harvey, P. H. 1994 Extinction rates can be  
571 estimated from molecular phylogenies. *Philosophical Transactions of the Royal Society*  
572 *of London Series B-Biological Sciences* **344**, 77-82.

573

574 Oksanen, J., Kindt, R., Legendre, P., O'Hara, B., Simpson, G. L., Solymos, P., Henry, M.  
575 H., Wagner, & S., Wagner, H. 2008 vegan: Community Ecology Package. R package  
576 version 1.15-0. <http://cran.r-project.org/>, <http://vegan.r-forge.r-project.org/>

577

578 Pagel, M. 1999 Inferring the historical patterns of biological evolution. *Nature* **401**, 877-  
579 884.

580

581 Pagel, M. & Lutzoni, F. 2002 Accounting for phylogenetic uncertainty in comparative  
582 studies of evolution and adaptation. In *Biological Evolution and Statistical Physics* (ed.  
583 M. Lässig & A. Valleriani), pp. 148-161. Berlin: Springer-Verlag.

584

585 Pagel, M. & Meade, A. 2007. BayesTraits ([www.evolution.rdg.ac.uk](http://www.evolution.rdg.ac.uk)). Version 1.0.  
586 Reading, UK.

587

588 Paradis E., Claude, J., & Strimmer, K. 2004 APE: analyses of phylogenetics and  
589 evolution in R language. *Bioinformatics* **20**, 289-290.  
590

591 Paradis, E. 2004 Can extinction rates be estimated without fossils? *Journal of Theoretical*  
592 *Biology* **229**, 19-30.  
593

594 Paradis, E. 2005 Statistical analysis of diversification with species traits. *Evolution* **59**, 1-  
595 12.  
596

597 Purvis, A. 2008 Phylogenetic approaches to the study of extinction. *Annual Review of*  
598 *Ecology, Evolution, and Systematics* **39**, 301-319.  
599

600 Purvis, A., Cardillo, M., Grenyer, R. & Collen, B. 2005 Correlates of extinction risk:  
601 phylogeny, biology, threat and scale. In *Conservation Biology 8: Phylogeny and*  
602 *Conservation* (ed. A. Purvis, J. L. Gittleman & T. Brooks). New York: The Zoological  
603 Society of London. pp 295-316.  
604

605 Purvis, A., Gittleman, J. L., Cowlshaw, G. & Mace, G. M. 2000 Predicting extinction  
606 risk in declining species. *Proceedings of the Royal Society B-Biological Sciences* **267**,  
607 1947-1952.  
608

609 Purvis, A., Nee, S. & Harvey, P. H. 1995 Macroevolutionary Inferences from Primate  
 610 Phylogeny. *Proceedings of the Royal Society of London Series B-Biological Sciences*  
 611 **260**, 329-333.  
 612  
 613 R Development Core Team. 2009 R: A language and environment for statistical  
 614 computing. R Foundation for Statistical Computing, Vienna, Austria.  
 615 ISBN 3-900051-07-0, URL <http://www.R-project.org>.  
 616  
 617 Rabosky, D. L. 2010 Extinction rates should not be estimated from molecular  
 618 phylogenies. *Evolution* (**early view**), 1-9.  
 619  
 620 Ray, D. A., Xing, J., Salem, A. H. & Batzer, M. A. 2006 SINEs of a nearly perfect  
 621 character. *Systematic Biology* **55**, 928-935.  
 622  
 623 Ray, D. A. & Batzer, M. A. 2005 Tracking Alu evolution in New World primates. *Bmc*  
 624 *Evolutionary Biology* **5**, -.  
 625  
 626 Rohlf, F. J. 2006 A comment on phylogenetic correction. *Evolution* **60**, 1509-1515.  
 627  
 628 Ronquist, F. & Huelsenbeck, J. P. 2003 MrBayes 3: Bayesian phylogenetic inference  
 629 under mixed models. *Bioinformatics* **19**, 1572-1574.  
 630

631 Roos, C., Schmitz, J. & Zischler, H. 2004 Primate jumping genes elucidate strepsirrhine  
632 phylogeny. *Proceedings of the National Academy of Sciences of the United States of*  
633 *America* **101**, 10650-10654.

634

635 Salem, A. H., Ray, D. A., Xing, J., Callinan, P. A., Myers, J. S., Hedges, D. J., Garber, R.  
636 K., Witherspoon, D. J., Jorde, L. B. & Batzer, M. A. 2003 Alu elements and hominid  
637 phylogenetics. *Proceedings of the National Academy of Sciences of the United States of*  
638 *America* **100**, 12787-12791.

639

640 Sanderson, D. S. 2002 *Estimating absolute rates of molecular evolution and divergence*  
641 *times: a penalized likelihood approach. Molecular Biology and Evolution* **19**, 101-109.

642

643 Schmitz, J., Ohme, M. & Zischler, H. 2001 SINE insertions in cladistic analyses and the  
644 phylogenetic affiliations of *Tarsius bancanus* to other primates. *Genetics* **157**, 777-784.

645

646 Seiffert, E. R., Simons, E. L. & Attia, Y. 2003 Fossil evidence for an ancient divergence  
647 of lorises and galagos. *Nature* **422**, 421-424.

648

649 Siegel, S. 1956 *Nonparametric Statistics for the Behavioral Sciences*. pp 23-26. New  
650 York: McGraw-Hill Book Company, Inc.

651

652 Smith, R. J. & Jungers, W. L. 1997 Body mass in comparative primatology. *Journal of*  
653 *Human Evolution* **32**, 523-559.

654

655 Stanley, S. M. 1973 Explanation for Copes Rule. *Evolution* **27**, 1-26.

656

657 Stevens, S. S. 1946 On the theory of scales of measurement. *Science* 103, 677-680.

658

659 Thorne, J. L., Kishino, H. & Painter, I. S. 1998 Estimating the rate of evolution of the rate  
660 of molecular evolution. *Molecular Biology and Evolution* **15**, 1647-1657.

661

662 Thorne, J. L. & Kishino, H. 2002 Divergence time and evolutionary rate estimation with  
663 multilocus data. *Systematic Biology* **51**, 689-702.

664

665 Vrba, E. S. 1984 What Is Species Selection. *Systematic Zoology* **33**, 318-328.

666

667 Webb, C. O., Ackerly, D. D., McPeck, M. A. & Donoghue, M. J. 2002. Phylogenies and  
668 community ecology. *Annu. Rev. Ecol. Syst.* **33**, 475-505

669

670 Wilson, D. E. & Reeder, D. M. 2005 *Mammal Species of the World*: Johns Hopkins  
671 University Press.

672

673 Xing, J., Wang, H., Han, K. D., Ray, D. A., Huang, C. H., Chemnick, L. G., Stewart, C.  
674 B., Disotell, T. R., Ryder, O. A. & Batzer, M. A. 2005 A mobile element based  
675 phylogeny of Old World monkeys. *Molecular Phylogenetics and Evolution* **37**, 872-880.

676



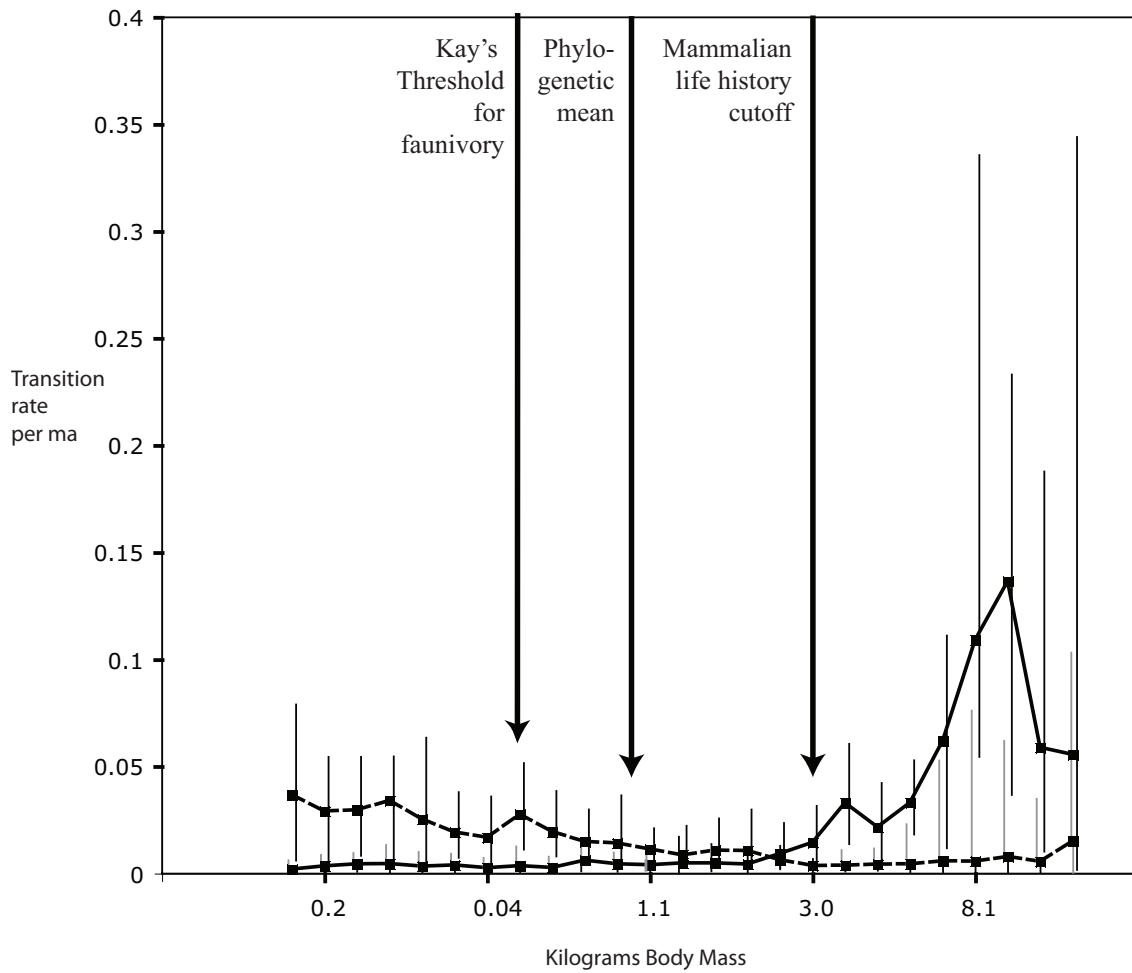
677 Xing, J. C., Witherspoon, D. J., Ray, D. A., Batzer, M. A. & Jorde, L. B. 2007 Mobile  
 678 DNA elements in primate and human evolution. *American Journal of Physical*  
 679 *Anthropology*, 2-19.  
 680  
 681 Yang, Z. H. & Yoder, A. D. 2003 Comparison of likelihood and Bayesian methods for  
 682 estimating divergence times using multiple gene loci and calibration points, with  
 683 application to a radiation of cute-looking mouse lemur species. *Systematic Biology* **52**,  
 684 705-716.  
 685  
 686 Yoder, A. D. & Yang, Z. H. 2004 Divergence dates for Malagasy lemurs estimated from  
 687 multiple gene loci: geological and evolutionary context. *Molecular Ecology* **13**, 757-773.  
 688  
 689

## Figure Legend

Figure 1. Sliding break point analysis of body mass dependent speciation and extinction rates on the consensus tree. Solid lines connecting points show rates for species that have larger body mass than the given point on the body mass X-axis (black 95% credible interval bars), while dashed lines are for species with smaller body mass (grey 95% credible interval bars). Triangles in Fig. 1a show inferred extinction rates and squares in Fig. 1b show inferred speciation rates. Thick vertical arrows correspond to our three *a priori* hypotheses.

Table 1. Fossil calibration ranges used to date molecular phylogenies

MRCA node	Min. Age (ma)	Max. Age (ma)	Source
<i>Homo-Pan</i>	5	8	Haile-Selassie (2001) Senut <i>et al.</i> (2001) Vignaud <i>et al.</i> (2002) Brunet <i>et al.</i> (2002)
<i>Homo-Pongo</i>	12.5	18	Kelley (2002)
<i>Papio-Theropithecus</i>	3.5	6.5	Leakey (1993)
extant Catarrhini	21.0	30.0	Young & MacLatchy (2004) Benefit & McCrossin (2002)
<i>Cebus-Saimiri</i>	12.5	NA	Hartwig & Meldrum (2002)
<i>Loris-Galago</i>	38	42	Seiffert <i>et al.</i> (2003)



**Supp. Fig. 1.** Sliding break point analysis of body mass dependent transition rates on the consensus tree. Solid lines connecting points show rates for species that have larger body mass than the given point on the body mass X-axis (black 95% credible interval bars), while dashed lines are for species with smaller body mass (grey 95% credible interval bars). Thick vertical arrows correspond to our three a priori hypotheses.